

# Implications of the Putnam Software Equation on Confidence Levels for Project Success

R. C. Tausworthe  
DSN Data Systems Section

*This article investigates the implications of assumed powerlaw relationships among size, duration, and effort on the probability that a given software project will be completed within its estimated schedule and manpower resources. Specifically, software development tasks are treated as sample points in a probability space characterized by three random variables: size, duration, and resource expenditure. The completion confidence factor is then computed. The most astonishing conclusion is the low confidence factor of the average project, significantly less than 25%. This low confidence factor is the result of correlation of the project duration and work effort by a tradeoff relationship referred to as "Putnam's software equation."*

## I. Introduction

The early-on estimation of the resources and schedule required for the development and maintenance of software has resulted in several resource and schedule models that accept such inputs as the enormity of the task, the physical, environmental, human, and management constraints assumed or known to be in effect, the history base of similar and dissimilar experience, and the means, alternatives, and technology available to the task. Such models attempt to predict the performance of humans doing the work, the events and expenditures in the development process, and the characteristics of the resulting product. The goal of such models has been primarily to predict the average characteristics of an envisioned project in such a way as to suffice for planning purposes.

The prediction of human task group behavior, however, may be viewed as a problem in estimating events in a stochas-

tic process governed by an unknown, and probably nonstationary, probability function. An optimum model can predict only to the limit imposed by the statistical characterization of the human activity.

In addition to the mean behavior of software projects, the statistical treatment can also lead to estimators for project parameters that permit the evaluation of certain risk factors.

The optimal cost prediction model, however, would require the precise quantification of all technical, environmental, and human-behavior parameters, and would combine these into a mathematical formula producing maximal likelihood or minimal variance results. Lacking this precise quantification, the best that one may hope for in a cost model is that it accommodate the principal factors affecting the estimate variance (or project risk).

There are a number of software cost models in existence, fourteen of which are summarized in Ref. 1. Most of these are least-square-error fits using power-law relationships among size, resource, and schedule parameters.

An IBM study (Ref. 2) reported the analysis of 60 software projects with respect to 68 situational variables believed to influence productivity. Of these, 29 showed a significant, high correlation with productivity, and were included in their estimation model. This model utilized power-law relationships for both the effects of the situational variables as well as the size, resource, and schedule variables.

Nine such models were evaluated for Air Force use as reported in Ref. 3. Model accuracy was measured and found to be best whenever a particular model was calibrated using representative historical data. Calibration was found to have greater effect on estimating accuracy than the precise model form. Conclusions based on statistics from the University of Maryland (Ref. 4) tended to confirm this hypothesis on power-law models.

All of the models above tended to focus on fitting measured statistics to relationships among pairs of parameters: effort vs size, duration vs size, effort vs duration, and average staff vs effort. Another model, the Rayleigh-Norden-Putnam model (Refs. 5 and 6), presupposes a power-law model among size, effort, and duration, a model calibrated using available industry data; however, the trivariate data for this calibration has not been published to the author's knowledge.

This article investigates the implications of assumed power-law relationships among size, duration, and effort on the probability that a given software project will be completed within its estimated schedule and manpower resources. Specifically, the paper treats software development tasks as sample points in a probability space characterized by three random variables: size, measured in delivered lines of source code, duration, measured in months, and resource expenditure, measured in man-months. Power-law equations are used to describe the relationships between expected values of the software random variables, and log-normal probability functions are used to approximate marginal and conditional distributions.

The paper then computes the *completion confidence factor*, defined as

$$C(T, W) = P\{t \leq T, w \leq W\}$$

i.e., the probability that the project duration,  $t$ , and work effort,  $w$ , will not simultaneously exceed values  $W$  and  $T$ , respectively.

## II. Software Project Parameter Relationships

We shall suppose, for the purposes of this article, that a software development can be characterized by its final delivered size, the total effort expended, and the overall length of time required. We shall denote these quantities as

$L$  = the number of kiloLines of delivered source code

$w$  = the work effort required in man-months

$t$  = the time duration in months

However, we shall find it more convenient to work with the logarithms of these:

$\log(L)$

$\log(w)$

$\log(t)$

In this way, power-law relationships among  $L$ ,  $w$ , and  $t$  become linear relationships in their logarithms.

For a given program to be written, an infinite ensemble of projects would not all produce the same values of  $t$ ,  $w$ , and  $L$ . Rather, one would observe a statistical distribution over the three-dimensional space spanned by these parameters. We shall thus treat software development characteristics as a probability space characterized by random variables  $\log(L)$ ,  $\log(w)$ , and  $\log(t)$ , governed by a probability density function,

$$p(\log(t), \log(w), \log(L))$$

In terms of the usual expected-value operation,  $E\{\cdot\}$ , the point

$$(\tau_0, \omega_0, \lambda_0) = (E\{\log(t)\}, E\{\log(w)\}, E\{\log(L)\})$$

represents the characteristics of the *average project* across the hypothetical ensemble of projects writing the particular hypothetical program.

### A. Effort-Duration Tradeoff Characteristics

The average-project point  $(\tau_0, \omega_0, \lambda_0)$  may be assumed to conform to a tradeoff law of the form

$$\tau_0 = g(\omega_0, \lambda_0)$$

That is, the average time required is influenced by the average effort applied and the average size of the task.

Let us therefore define zero-mean random variables:

$$\tau = \log(t) - \tau_0$$

$$\omega = \log(w) - \omega_0$$

$$\lambda = \log(L) - \lambda_0$$

having standard deviations  $\sigma_\tau$ ,  $\sigma_\omega$ , and  $\sigma_\lambda$ , respectively, and probability density function  $p(\tau, \omega, \lambda)$ . We seek to approximate the probability (confidence) that a project will be completed within schedule *and* manpower resources,

$$\begin{aligned} C(T, W) &= P\{t \leq T, w \leq W\} \\ &= P\{\tau \leq X, \omega \leq Y\} \\ &= \int_{-\infty}^X \int_{-\infty}^Y \int_{-\infty}^{+\infty} p(\tau, \omega, \lambda) d\lambda d\omega d\tau \\ &= \int_{-\infty}^X \int_{-\infty}^Y p(\tau, \omega) d\omega d\tau \end{aligned}$$

where

$$X = \log(T) - \tau_0$$

$$Y = \log(W) - \omega_0$$

The values  $X$  and  $Y$  are logarithmic schedule and effort margins measured from the average-project point.  $X$  and  $Y$  are related via the tradeoff law as follows:

$$X = \log(T) - g(\log(W) - Y, \lambda_0)$$

It is clear that  $C(T, W)$  depends on the average-project point, i.e., on the way expected effort has been traded with expected duration for a certain expected size. We shall optimize this situation by choosing the tradeoff to maximize the confidence,

$$C_0(T, W) = \max_{\tau_0 = g(\omega_0, \lambda_0)} C(T, W)$$

The condition for maximization of confidence is

$$\frac{dC(T, W)}{d\omega_0} = 0$$

which yields

$$\frac{dX}{dY} \int_{-\infty}^Y p(X, \omega) d\omega + \int_{-\infty}^X p(\tau, Y) d\tau = 0$$

Note that  $X$  may be eliminated from this equation by using the relationship imposed by the tradeoff law above. Thence the equation can be solved for  $Y$ , and then  $X$  and  $(\tau_0, \omega_0)$  found. These margin values will be denoted  $X_0$  and  $Y_0$ , respectively.

Note also that the solution above is the same as if  $C(T, W)$  were maximized with respect to  $Y$ , subject to the constraint imposed by the tradeoff law.

## B. Putnam's Software Equation

Putnam (Ref. 6) has postulated a tradeoff between size, effort, and duration, and has given it the form

$$L = c_k w^p t^q$$

or

$$\log(L) = \eta + p \log(w) + q \log(t)$$

or, if we presume that the tradeoff is to be valid in a neighborhood of the average-project point,

$$\lambda = p\omega + q\tau$$

with

$$\lambda_0 = \eta + p\omega_0 + q\tau_0$$

The function  $g(\ )$  of the previous section is thus linear:

$$\tau_0 = g(\omega_0, \lambda_0) = -\eta/q + \lambda_0/q - p\omega_0/q$$

and  $X$  and  $Y$  are related by

$$qX + pY = \eta + p \log(W) + q \log(T) - \lambda_0 = \Delta$$

The parameter  $\Delta$  defined by this equation may be viewed as a *size margin*, because  $\Omega = \eta + p \log(W) + q \log(T)$  is the virtual expected size when duration  $T$  and effort  $W$  are expended, and  $\lambda_0$  is the expected size of the project. Note: it is not necessary that the tradeoff equation apply at  $(T, W)$ .  $\Omega$  is merely the constant defined by  $T$  and  $W$  in the tradeoff rule above.

Putnam calls the tradeoff law the "software equation," and specifically postulates the values

$$p = 1/3$$

$$q = 4/3$$

for a *tradeoff exponent ratio* value of  $r = q/p = 4$ . These values were obtained by Putnam after a study of several large software implementation projects.

The author has been unable so far to locate collaborative published statistics in the open literature supporting the  $r = 4$  figure. However, in private communications, the author has been informed by individuals having access to software productivity data bases that this particular statistic would not be difficult to compute, when undertaken.

The particular value  $r = 4$  implies that a factor of 16 times as much manpower is required to shorten a schedule by a factor of 2. However, it well may be that  $r$  depends on project size. A value  $r = 4$  may adequately describe localized changes in project parameters by a few percent about the average-project point for very large undertakings, but it may not apply to small-to-medium-sized efforts, where one may truly halve the overall duration by application of increased resources. For medium-scale efforts, it has been suggested that perhaps only an increase of 1.5 times the man-months effort is needed when the duration is halved, or  $r = \log_2(1.5) = 0.484$ . We shall discuss the influence of the value  $r$  on the confidence factor at greater length later in this article.

We shall therefore keep the values of the software equation coefficients undetermined for the present in order to be a bit more general in our approach, and later we shall solve for these, to the extent possible, using published empirical data.

### C. Derivation of the Software Equation

If one were given a set of point-data  $\{(\tau, \omega, \lambda)\}$  taken from an ensemble of projects of various durations, efforts, and sizes, then one could perform a least-squared-error fit to the data to determine general relationships among the random variables. Specifically, for the duration variable  $\tau$  the best-fit curve would be (Ref. 7):

$$\tau_{\omega, \lambda} = E\{\tau|\omega, \lambda\}$$

We shall assume, in keeping with the Putnam model, that this relationship is approximately a linear one

$$\tau_{\omega, \lambda} = (\lambda - p\omega)/q$$

or

$$\lambda = p\omega + q\tau_{\omega, \lambda}$$

for appropriately chosen  $p$  and  $q$ . We have purposely chosen these coefficients to correspond to those appearing in the Putnam equation for later discussion.

The orthogonality principle in probability theory (Ref. 7) states that the error between  $\tau$  and  $\tau_{\omega, \lambda}$  is uncorrelated with both  $\omega$  and  $\lambda$ . This provides two equations

$$E\{\omega\lambda\} = p\sigma_{\omega}^2 + qE\{\omega\tau\}$$

$$\sigma_{\lambda}^2 = pE\{\omega\lambda\} + qE\{\lambda\tau\}$$

In a similar fashion, one may seek to find relationships among each pair of random variables by finding the least-squared-error functions

$$\omega_{\lambda} = E\{\omega|\lambda\}$$

$$\tau_{\lambda} = E\{\tau|\lambda\}$$

$$\tau_{\omega} = E\{\tau|\omega\}$$

Here there is strong evidence (Refs. 2, 4, and 8) that a linear approximation to each of these functions is valid, say

$$\omega_{\lambda} = E\{\omega|\lambda\} = \alpha\lambda$$

$$\tau_{\lambda} = E\{\tau|\lambda\} = \beta\lambda$$

$$\tau_{\omega} = E\{\tau|\omega\} = \gamma\omega$$

Figures 1, 2, and 3 show scatter diagrams in support of these hypotheses.

Application of the orthogonality principle to these expectations produces the relationships

$$E\{\omega\lambda\} = \alpha\sigma_{\lambda}^2$$

$$E\{\lambda\tau\} = \beta\sigma_{\lambda}^2$$

$$E\{\omega\tau\} = \gamma\sigma_{\omega}^2$$

which may be substituted into the two equations found earlier to obtain

$$\alpha p + \beta q = 1$$

$$\alpha p + \alpha \gamma q = \alpha^2 \sigma_{\lambda}^2 / \sigma_{\omega}^2$$

provided that neither  $\sigma_{\lambda}$  nor  $\sigma_{\omega}$  is zero.

Normally, one could solve these two simultaneous equations for  $p$  and  $q$  directly, whereupon a comparison with the Putnam values would be immediate, in terms of published best-fit parameters. However, we expect that  $\beta$  is approximately equal to  $\alpha\gamma$ , for the following reason: Note that  $\alpha$  is the best-fit coefficient relating  $\lambda$  to  $\omega$ , and  $\gamma$  is the best-fit coefficient relating  $\omega$  to  $\tau$ . Hence their product should be the coefficient relating  $\lambda$  to  $\tau$ , and  $\beta$  is the actual best-fit coefficient relating  $\lambda$  to  $\tau$ . If equality were the case, the two equations above would be singular, so no unique solution for  $p$  and  $q$  would then result.

But it is not the case that  $\beta = \alpha\gamma$ . If it were, the righthand sides of the two equations would have to be equal, a condition that would require

$$\alpha^2 = \sigma_\omega^2 / \sigma_\lambda^2 = \alpha^2 + \sigma_{\omega|\lambda}^2 / \sigma_\lambda^2$$

(we derive this latter relationship later), leading to a contradiction.

It is the case, however, that  $\beta$  is near enough to  $\alpha\gamma$  that small errors are greatly magnified in the solutions for  $p$  and  $q$ . For this reason, it will be necessary to find  $p$  and  $q$  by other means, such as an actual two-parameter linear least-squares-fit of the raw data.

We can nevertheless express the values of  $p$  and  $q$  in terms of  $r = q/p$ ,

$$p = \frac{1}{\alpha + \gamma r}$$

$$q = \frac{r}{\alpha + \gamma r}$$

In summary, we have assumed thus far only the following: (1) that the function  $E\{\tau|\omega, \lambda\}$ , which minimizes the least-squared-error, is linear in both  $\omega$  and  $\tau$ , and (2) that each of the best-fit functions  $E\{\omega|\lambda\}$ ,  $E\{\tau|\lambda\}$ , and  $E\{\tau|\omega\}$  are also linear in their variables. These assumptions are all certainly approximately true, as supported by published analyses.

#### D. Program Size Statistics

For a given program to be written, one may estimate the mean and variance in size of the program by a number of techniques. Let us suppose that the method chosen produces representative estimates, and let us approximate the distribution of actual lines of code finally produced by a normal probability function,

$$p(\lambda) = Z(\lambda/\sigma_\lambda)$$

where (Ref. 9)

$$Z(x) = (2\pi)^{-0.5} \exp(-x^2/2)$$

Program size statistics are often postulated to fit the beta distribution rather than the normal, so as to avoid the theoretical occurrence of extremely large values of  $\lambda$ , which are permitted in the normal model (with low probabilities), but are absent in actuality. We shall discuss the implications of this later in the section on accuracy considerations.

#### E. Joint Effort-Duration Statistics

Let us now compute the function  $p(\tau, \omega)$ , which will permit us to evaluate the confidence integral for a particular project ensemble.

We shall assume that, if  $\omega$  and  $\lambda$  are given, the density of the remaining random variable,  $\tau$ , can be approximated by a normal density, in which the conditional mean in  $\tau$  is determined by the software equation. Similarly, given  $\lambda$ , the density of  $\omega$  can be estimated by the normal density, with its conditional mean as that value determined earlier. That is, we assume

$$p(\tau|\omega, \lambda) = Z((\tau - \tau_{\omega, \lambda})/\sigma_{\tau|\omega, \lambda})$$

$$p(\omega|\lambda) = Z((\omega - \omega_\lambda)/\sigma_{\omega|\lambda})$$

where  $\sigma_{\tau|\omega, \lambda}$  is the standard deviation of  $\tau$  given both  $\omega$  and  $\lambda$ , and  $\sigma_{\omega|\lambda}$  is the deviation of  $\omega$  given  $\lambda$ . Scatter diagrams of  $\omega$  vs  $\lambda$  published in the literature, such as that shown in Fig. 1, indicate that the  $p(\omega|\lambda)$  approximation is reasonable. Further, such studies provide measured values for  $\sigma_{\omega|\lambda}$ .

The sought-for joint density is then

$$\begin{aligned} p(\tau, \omega) &= \int_{-\infty}^{+\infty} p(\tau|\omega, \lambda) p(\omega|\lambda) p(\lambda) d\lambda \\ &= \delta Z(\tau/\sigma_\tau) Z(\delta(\omega/\sigma_\omega - \rho\tau/\sigma_\tau)) \end{aligned}$$

This is a joint normal density in which the standard deviations  $\sigma_\tau$  and  $\sigma_\omega$ , correlation coefficient  $\rho$ , and parameter  $\delta$  are given by

$$\sigma_\tau^2 = \sigma_{\tau|\omega, \lambda}^2 + \frac{\sigma_{\omega|\lambda}^2 (\sigma_{\omega|\lambda}^2 + \sigma_\lambda^2)}{r^2 \sigma_\omega^2}$$

$$\sigma_\omega^2 = \sigma_{\omega|\lambda}^2 + \alpha^2 \sigma_\lambda^2$$

$$\rho = -\sigma_{\omega|\lambda}^2 / r \sigma_{\tau} \sigma_{\omega}$$

$$\delta = (1 - \rho^2)^{-0.5}$$

(the equation above for  $\sigma_{\omega}^2$  is that referred to earlier, which yielded the contradiction in the equality  $\beta = \alpha\gamma$ ).

Even though  $\sigma_{\tau|\omega,\lambda}$  has not been estimated in the literature, the value of  $\sigma_{\tau}$  can be evaluated, as will be shown later, to be

$$\sigma_{\tau}^2 = \sigma_{\tau|\lambda}^2 + \beta^2 \sigma_{\lambda}^2$$

All parameters needed to evaluate the joint density are therefore available from published data.

Note, however, that since  $|\rho| \leq 1$  and since  $r$  is unaffected by the value of  $\sigma_{\lambda}$ , the value of  $r$  is constrained by the lower bound

$$r \geq \frac{\sigma_{\omega|\lambda}}{\sigma_{\tau|\lambda}}$$

This relationship constrains  $p$  and  $q$  as follows:

$$p \leq \frac{\sigma_{\tau|\lambda}}{\alpha \sigma_{\tau|\lambda} + \gamma \sigma_{\omega|\lambda}}$$

$$q \geq \frac{\sigma_{\omega|\lambda}}{\alpha \sigma_{\tau|\lambda} + \gamma \sigma_{\omega|\lambda}}$$

## F. Effort Statistics

Integrating the joint density over  $\tau$  yields the approximate marginal density of  $\omega$ ,

$$p(\omega) = Z(\omega/\sigma_{\omega})$$

in which  $\sigma_{\omega}$  is given above.

## G. Duration Statistics

Integration of the joint density above with respect to  $\omega$  yields a normal approximation for  $p(\tau)$  whose variance has a term  $\sigma_{\tau|\omega,\lambda}^2$ . However, the variance may be computed another way: the published statistics of  $\tau$  vs  $\lambda$  indicate that a normal approximation is appropriate:

$$p(\tau|\lambda) = Z((\tau - \tau_{\lambda})/\sigma_{\tau|\lambda})$$

where  $\sigma_{\tau|\lambda}$  is the standard deviation in  $\tau$ , given  $\lambda$ , obtained from the published data. Averaging over  $\lambda$  produces the normal marginal density for  $\tau$ ,

$$p(\tau) = Z(\tau/\sigma_{\tau})$$

in which the variance of  $\tau$  is, as stated earlier,

$$\sigma_{\tau}^2 = \sigma_{\tau|\lambda}^2 + \beta^2 \sigma_{\lambda}^2$$

Thus, the deviation in duration can be calculated directly from the measured and estimated deviations in  $\omega$  and  $\lambda$ , respectively. Hence, the need to compute  $\sigma_{\tau|\omega,\lambda}^2$  is removed.

## H. Computation of Confidence Factor

Integration of the joint normal density  $p(\tau, \omega)$  to produce  $C(T, W)$  as indicated earlier yields a known (but untabulated) function (Ref. 9)

$$C(T, W) = L(-X/\sigma_{\tau}, -Y/\sigma_{\omega}, \rho)$$

where  $L(x, y, \rho)$  is defined as the double integral

$$L(x, y, \rho) = \int_x^{\infty} \int_y^{\infty} \delta Z(u) Z(\delta(v - \rho u)) du dv$$

Values for  $L(x, y, \rho)$  may be found by numerical integration.

One particular case of interest can be evaluated directly, however, namely the confidence factor for the average project,  $C(T_0, W_0)$ . This results when  $X = Y = 0$ , giving

$$C(T_0, W_0) = L(0, 0, \rho)$$

$$= \frac{1}{4} + \frac{\arcsin \rho}{2\pi}$$

Note that since  $\rho$  is a negative quantity, the *confidence in average project success is less than 25%*. To raise the confidence factor to reasonable levels, it is therefore necessary to increase the schedule and effort margins, i.e.,  $X$  and  $Y$ , to positive values.

Evaluation of the previous conditions for finding  $X_0$  and  $Y_0$  produces the equation

$$Z(aX_0) P(b - cX_0) = h Z(d - eX_0) P(f + gX_0)$$

where  $P(x) = (1 + \operatorname{erf}(x/2^{0.5}))/2$ , in which  $\operatorname{erf}(\cdot)$  is the well-known error function (Ref. 9), and the coefficients  $a$

through  $h$  are given by

$$a = 1/\sigma_\tau$$

$$b = d\delta$$

$$c = \delta(r/\sigma_\omega + \rho/\sigma_\tau)$$

$$d = \Delta/p\sigma_\omega$$

$$e = r/\sigma_\omega$$

$$f = -b\rho$$

$$g = c\rho + 1/\delta\sigma_\tau$$

$$h = e\sigma_\tau$$

The solution  $X_0$  to this equation may be found numerically using Newton's method, and then the corresponding  $Y_0$  computed.

The confidence factor is then computed by evaluating

$$C_0(T, W) = L(-X_0/\sigma_\tau, -Y_0/\sigma_\omega, \rho)$$

It is interesting to note in the optimization equation when  $\Delta = 0$  that  $X_0 = Y_0 = 0$  is *not* a solution. Instead, one can show that  $X_0 < 0$  and  $Y_0 > 0$  for this situation. This means that the confidence will be somewhat greater for a project expected to last somewhat longer than  $\log(T)$  with somewhat less manpower than  $\log(W)$  than it will be for a project with expected duration equal to  $\log(T)$  and expected manpower equal to  $\log(W)$ !

### III. Evaluation

In this section we shall compute the confidence characteristics of projects taken from two sources of published data. These two sources were not chosen because of their similarities or differences, or because the data was particularly extensive or internally consistent. Rather, the sources were used because they contained enough information to compute the model parameters of interest.

#### A. Walston-Felix Data

Data published by Walston and Felix (Ref. 2) provide the following parameter values of interest:

$$\alpha = 0.91$$

$$\sigma_{\omega|\lambda} = 0.92$$

$$\beta = 0.36$$

$$\sigma_{\tau|\lambda} = 0.542$$

$$\gamma = 0.35$$

$$\sigma_{\tau|\omega} = 0.419$$

$$\beta - \alpha\gamma = 0.0415$$

Even though (Ref. 2) contains a means for approximately normalizing the effects of the 29 noted situational and environmental variables, the values above do not reflect this (partial) normalization.

We shall assume, purely for illustrative purposes, that a PERT (Ref. 10) estimation scheme is used to estimate  $\lambda$  and its variance, and that  $\lambda_{max} = \log(1.5\lambda_0)$  and  $\lambda_{min} = \log(\lambda_0/2)$ , so that we have

$$\sigma_\lambda = (\lambda_{max} - \lambda_{min})/6 = 0.183$$

We then obtain from the formulas above

$$\sigma_\omega = 0.935$$

$$\sigma_\tau = 0.546$$

The bounds on  $r$ ,  $p$ , and  $q$  are

$$r \geq 1.7$$

$$p \leq 0.665$$

$$q \geq 1.13$$

A range of confidence factor calculations for the average-project point is shown in Table 1.

The Putnam value  $p = 1/3$  would require  $r = 5.97$ , according to the formulas above, rather than the  $r = 4$  value used by Putnam. The  $r = 4$  value, on the other hand, together with the measured values above, produces  $p = 0.433$ . If  $q = 4/3$  were the case, then  $r$  would have to be 2.28.

The average-project confidence factors for these cases are shown in Fig. 4 as a function of  $r$ . The optimized confidence factor characteristics are plotted in Fig. 5 as a function of the size margin parameter  $\Delta$ . The optimum duration and effort margins are given in Fig. 6.

#### B. Freburger-Basili Data

Data published by Freburger and Basili (Ref. 4) give

$$\alpha = 0.986$$

$$\sigma_{\omega|\lambda} = 0.378$$

$$\beta = 0.203 \quad \sigma_{\tau|\lambda} = 0.315$$

$$\gamma = 0.210 \quad \sigma_{\tau|\omega} = 0.300$$

$$\gamma = 0.210$$

$$\beta - \alpha\gamma = -0.0004$$

The data from which these parameters were derived were taken from a more carefully controlled environment and situation than the Walston-Felix data was. Computations using these values result in

$$\sigma_{\omega} = 0.419$$

$$\sigma_{\tau} = 0.317$$

The bounds on  $r$ ,  $p$ , and  $q$  are

$$r \geq 1.202$$

$$p \leq 0.807$$

$$q \geq 0.971$$

The values of the correlation coefficient and average-project confidence factor deriving from these values are shown in Table 2.

A value of  $r = 4$  implies  $p = 0.547$ , rather than the Putnam  $1/3$ , and the  $p = 1/3$  value implies  $r = 9.59$ . A  $q$ -value of  $4/3$  requires  $r = 1.82$ .

The average-project confidence factors above are plotted in Fig. 4 vs  $r$  for comparison with the Walston-Felix data. The optimized confidence factor behavior as a function of the size margin is shown in Fig. 7, and the optimum duration and effort margins are given in Fig. 8.

### C. Comparison of Results

The Freburger-Basili data generally show a higher confidence factor than do the Walston-Felix data, both for the average project and the optimized project. The difference may be due to the extreme care and consistency taken in recording the data in the former case. The latter data were taken from a diverse set of projects in a wide variation of environments and situational factors, whereas the former were taken under very controlled and recorded conditions. If both sets of data were normalized to remove situational effects, as described in the Walston-Felix paper, then perhaps the agreement would be closer.

Nothing can be learned about the value of  $r$  from either of these data sets, as published. The variation of  $X_0$  and  $Y_0$  for the two cases is radically different.

## IV. Accuracy Considerations

Accuracy in the figures produced by the model above depends not only on the assumptions about normality of the logarithms and linearity of the conditional expectations, but also on the accuracy of the inferences from statistical measures to the model parameter values. By assuming an effort-duration tradeoff law, we have admitted that the average project facing a given situation can be influenced by the expected allocation of effort and duration. Yet the published statistics derive from an ensemble of projects tasked with an ensemble of very different situations. The marked differences between the Walston-Felix data and the Freburger-Basili data are strong indicators that situational factors, such as technology, organization, experience, and environment, do influence project performance. Walston and Felix confirm this hypothesis in their article, where they find significant correlation between performance and 29 situational variables.

Therefore, the published statistics represent an averaging over the model we have presented here with respect to the ensemble of industry projects. We may thus expect that the use of measured variances taken from such data will exceed the variances assumed in the model above. In this respect, use of the Walston-Felix or Freburger-Basili numbers may produce somewhat pessimistic results, unless the situational factors can be normalized to some extent by an appropriate method.

Frequently, the statistics of a given bounded data set are presumed to be characterized by a beta density,

$$p(x) = A (x - x_1)^b (x_2 - x)^c$$

over the range  $x_1 \leq x \leq x_2$ , with  $p(x) = 0$  elsewhere.  $A$  is a constant chosen to make the density have unit area. Typically, values  $b = c = 2$  are used in the so-called PERT (Ref. 10) estimation technique.

This density has peak value  $x_0$  given by

$$x_0 = \frac{bx_2 + cx_1}{d}$$

where  $d = b + c$ . The mean and variance of the distribution are

$$\mu = \frac{x_1 + dx_0 + x_2}{d + 2}$$



$$\sigma^2 = \frac{(d + bc + 1)(x_2 - x_1)^2}{(d + 2)^2 (d + 3)}$$

The values of  $b$  and  $c$  for a given  $d$  may be adjusted to accommodate nonsymmetry in the distribution.

Let us now consider how well the normal probability density approximates the beta density. Let us compare both as zero-mean, unit-variance densities with respect to two measures: first, the root-sum-square error, and second, the probability that the normal variate exceeds the beta-distribution cutoff point. The conditions for the beta density having zero mean and unit variance are

$$x_1 = -(b + 1) \left[ \frac{d + 3}{bc + d + 1} \right]^{0.5}$$

$$x_2 = -\frac{(c + 1)}{(b + 1)} x_1$$

The first comparison function, the root-sum-error, is

$$\epsilon_1 = \left[ \int_{-\infty}^{+\infty} (p(x) - Z(x))^2 dx \right]^{0.5}$$

which can be found by numerical integration. The second comparison function can be evaluated directly,

$$\begin{aligned} \epsilon_2 &= [1 - P(-x_1)] + [1 - P(x_2)] \\ &= 0.5 [\operatorname{erfc}(-x_1/2^{0.5}) + \operatorname{erfc}(x_2/2^{0.5})] \end{aligned}$$

where  $\operatorname{erfc}(\cdot)$  is the complementary error function (Ref. 9).

The values of both  $\epsilon_1$  and  $\epsilon_2$  are shown in Table 3 below. From these one may note that both errors decrease with  $d$ , and for a given  $d$ , both are least when  $b$  and  $c$  are equal (i.e., when the density is symmetric, as is  $Z(x)$ ). We may therefore conclude that for a reasonable  $d$  (e.g.,  $d = 4$  in the PERT technique), the error introduced into the model by assuming a normal distribution, rather than a beta distribution, is negligible with respect to the much larger uncertainties involved with estimating the parameters required by the model.

## V. Summary and Conclusions

There is something unsettling, or at least disappointing, in the results of these analyses, that the average project will probably fail in meeting one of its performance goals. How-

ever, this likelihood is implied by the combined assumptions of the popular software cost estimators assembled into the model of this article. All of these assumptions individually sound reasonable, fit together in a consistent, logical way, and are backed up by measurements of one sort or another.

Moreover, the model predicts that a significant contingency bias in planned manpower and schedule will be required to reach acceptable engineering risk levels.

What may be even more unsettling is that this theory may apply to other kinds of projects as well, where there is a log-linear, manpower-schedule tradeoff possible. A study of productivity and schedule statistics in this area might be very revealing.

However, for the moment, we may consider that part of what has been called "the software crisis," i.e., that most projects seem to fail one way or another, may not be the fault of either the programmer or his management; for even if they were able to estimate exactly what the average project would do in their given situation, the odds are that they would still fail, if they planned for the average.

What this paper shows, nevertheless, is that *planning for an average project contributes to the crisis*. To succeed within performance goals, it is necessary to do the following things:

- (1) Estimate carefully the size, and bounds on size, of the task.
- (2) Negotiate manpower and schedule constraints.
- (3) Determine the risks associated with failure of the project in both manpower and schedule dimensions.
- (4) Negotiate an appropriate confidence factor under which the risks are acceptable.
- (5) Determine  $(T, W)$  that will produce the desired confidence within manpower and schedule constraints using the model above.
- (6) Schedule the task to utilize manpower  $W$  and duration  $T$ .
- (7) If a  $(T, W)$  cannot be found that is compatible with risk and constraints on manpower and schedule, renegotiate.

As mentioned earlier, the model will probably produce somewhat pessimistic results if published industry parameter values are used. Accuracy in estimating confidence factors and planned performance margins can come only through precalibration of the model parameters in the particular software development environment and normalization of the situational

factors that may influence that particular project in the given environment. The Thibodeau study (Ref. 3) also confirmed that such calibration is needed for more basic modeling accuracy considerations.

To promote better accuracy, it is necessary to develop a basic software cost model, such as that reported in Ref. 2 and extended by the author in Ref. 11, that will tend to normalize

the environmental and situational factors a particular project faces, to maintain an historical archive of measured project characteristics, and to analyze the collected data with respect to the model to extract the needed model parameters.

In particular, the value of  $r$  needs to be explicitly measured. The analyses of this paper are inconsistent with the Putnam  $p$ ,  $q$ , and  $q/p$ .

## References

1. *Quantitative Software Models*, Report SRR-1, Data and Analysis Center for Software, RADC, Griffiss Air Force Base, New York, AFB, N.Y., March 1979.
2. Walston, C.E., and Felix, C.P., "A Method of Programming Measurement and Estimation," *IBM System Journal*, Vol. 16, No. 1, 1977.
3. Thibodeau, R., *An Evaluation of Software Cost Estimating Models*, Report CR-1-940, prepared for Rome Air Development Center, Air Force Systems Command, Griffiss Air Force Base, New York, by General Research Corp., Huntsville, Ala., April 10, 1981.
4. Freburger, K., and Basili, V. R., *The Software Engineering Laboratory: Relationship Equations*, TR-764, SEL-3, NSG-5123, University of Maryland, College Park, Md., May 1979.
5. Norden, P. V., "Project Life Cycle Modeling," Software Life-Cycle Management Workshop, United States Army Computer Systems Command, August 1977, pp. 217-306.
6. Putnam, L. H., "Progress in Modeling the Software Life Cycle in a Phenomenological Way to Obtain Engineering Quality Estimates and Dynamic Control of the Process," in *Software Cost Estimating and Life-Cycle Control*, IEEE Computer Society Press, Los Alamitos, Ca., 1980, pp. 183-206.
7. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Book Co., N.Y., 1965, pp. 104, 147, 241-243, 250-258.
8. Nelson, R., *Software Data Collection and Analysis, (Partial Report)*, Rome Air Force Development Center, Data and Analysis Center for Software, Griffiss Air Force Base, New York, September 1978.
9. *Handbook of Mathematical Functions*, AMS-55, National Bureau of Standards, U. S. Government Printing Office, Washington D.C., 1964, pp. 297-298, 931-937.
10. The author is unable to trace the definitive reference for this work. The usage cited is exposed in Putnam, L. H., "Example of an Early Sizing, Cost and Schedule Estimate for an Application Software System," *COMPSAC '79*, November 13-16, 1978.
11. Tausworthe, R. C., *Deep Space Network Software Cost Estimation Model*, Publication 81-7, Jet Propulsion Laboratory, Pasadena, Ca., Apr. 15, 1981.

**Table 1. Average project confidence factors as a function of the tradeoff ratio  $r$ , from Walston-Felix data**

$r$	$p$	$q$	$\rho$	$C(T_0, W_0)$
2.00	0.621	1.24	-0.829	0.094
2.28	0.586	1.33	-0.729	0.120
4.00	0.433	1.73	-0.415	0.182
5.97	0.333	1.99	-0.277	0.205

**Table 2. Average project confidence factors as a function of the tradeoff ratio  $r$ , from Freburger-Basili data**

$r$	$p$	$q$	$\rho$	$C(T_0, W_0)$
1.83	0.730	1.33	-0.590	0.150
2.00	0.711	1.42	-0.539	0.159
4.00	0.547	2.19	-0.269	0.207
9.59	0.333	3.20	-0.112	0.232

**Table 3. Beta density vs normal density goodness of fit**

$b$	$c$	$d$	$\epsilon_1$	$\epsilon_2$
1	1	2	0.088	0.025
1	2	3	0.091	0.024
1	3	4	0.107	0.031
1	4	5	0.122	0.037
2	1	3	0.091	0.024
2	2	4	0.057	0.008
2	3	5	0.058	0.008
2	4	6	0.067	0.010
3	1	4	0.107	0.031
3	2	5	0.058	0.008
3	3	6	0.043	0.003
3	4	7	0.043	0.003
4	1	5	0.122	0.036
4	2	6	0.067	0.010
4	3	7	0.043	0.003
4	4	8	0.033	0.001

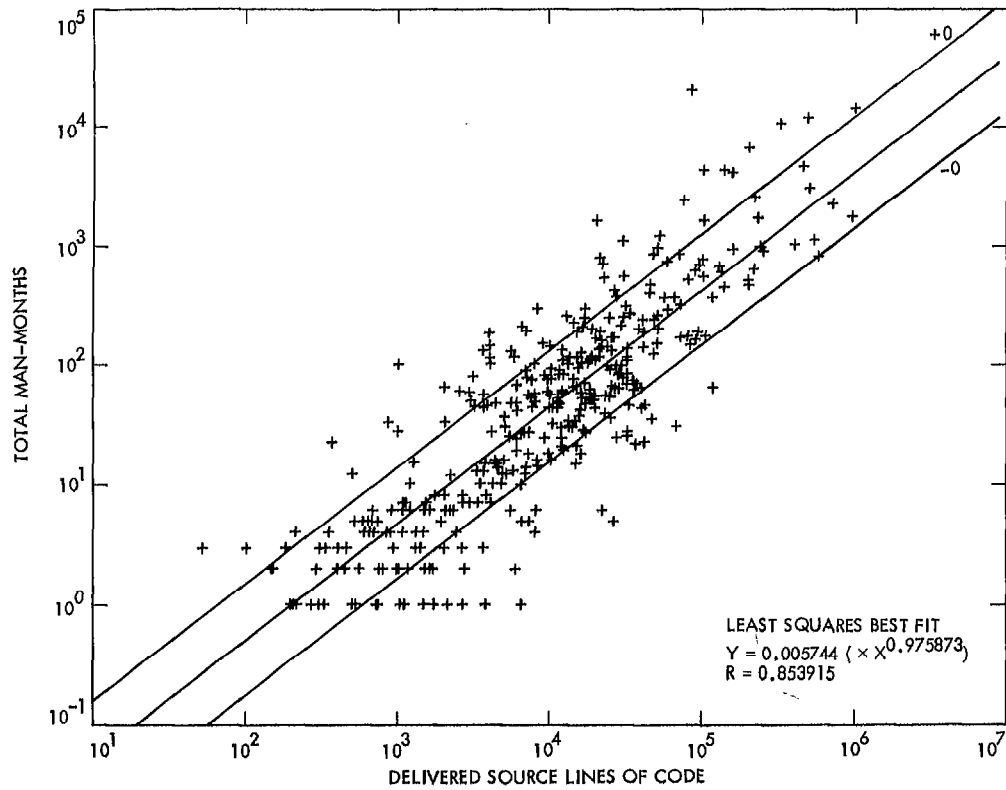


Fig. 1. RADC software project effort vs size data (Ref. 8)

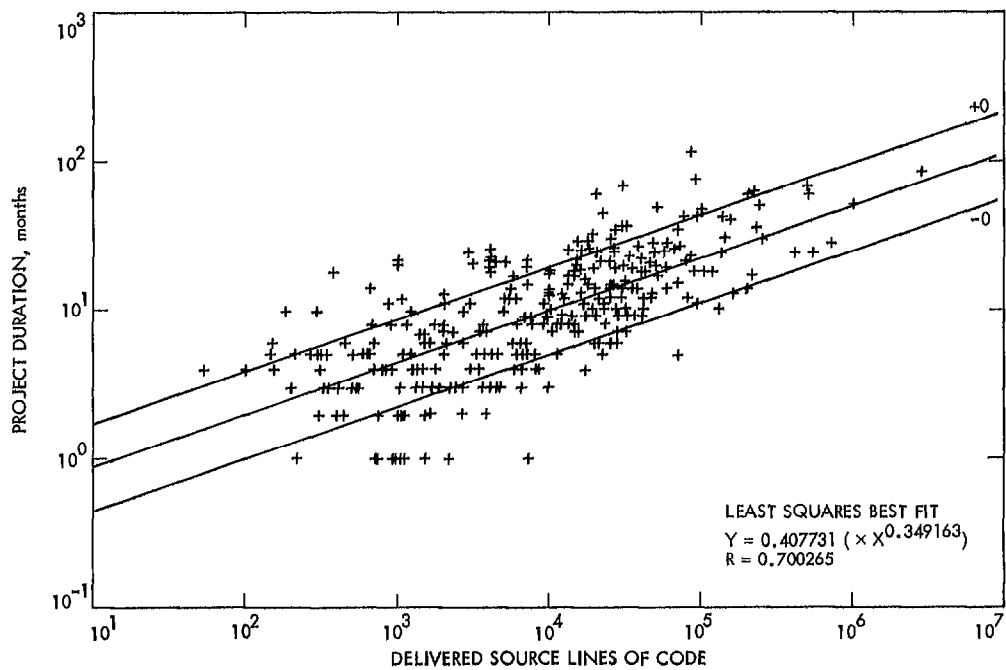


Fig. 2. RADC software project duration vs size data (Ref. 8)

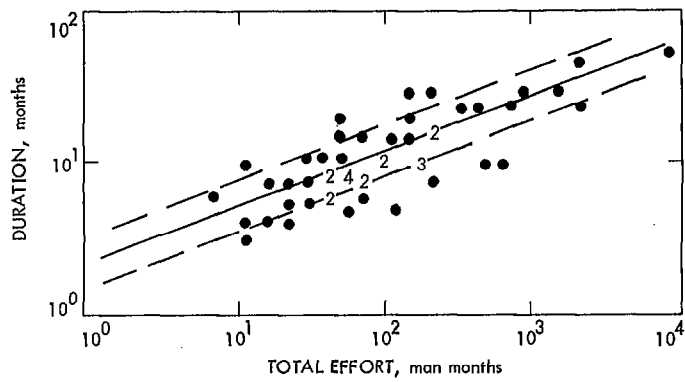


Fig. 3. IBM software project duration vs effort data (Ref. 2)

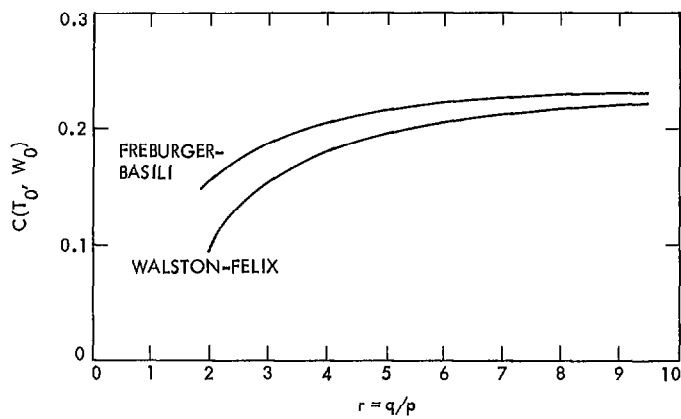


Fig. 4. Average-project confidence factors as a function of  $r$  for both the Walston-Felix data and Freiburg-Basili data

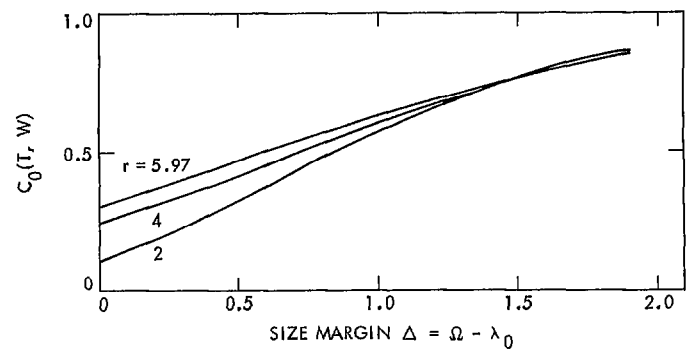


Fig. 5. Optimum project confidence factor for given size margin  $\Delta$ , Walston-Felix data

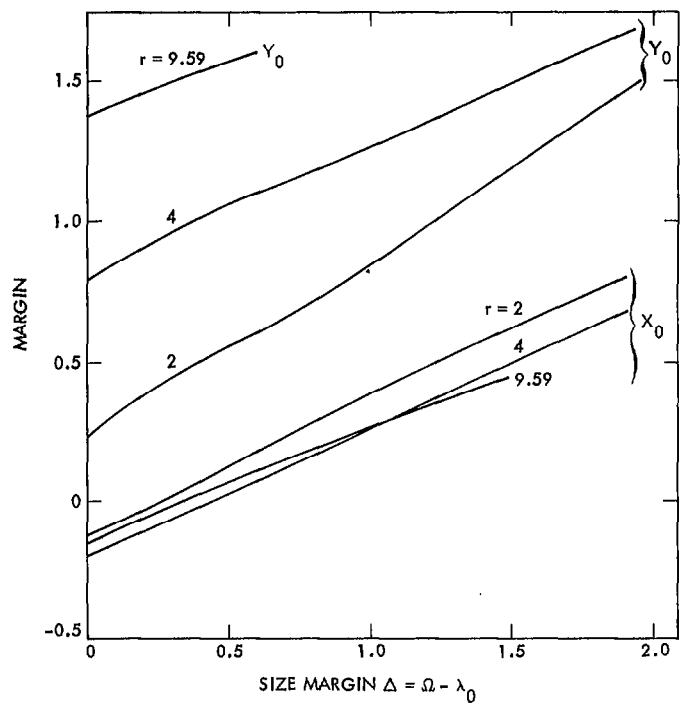
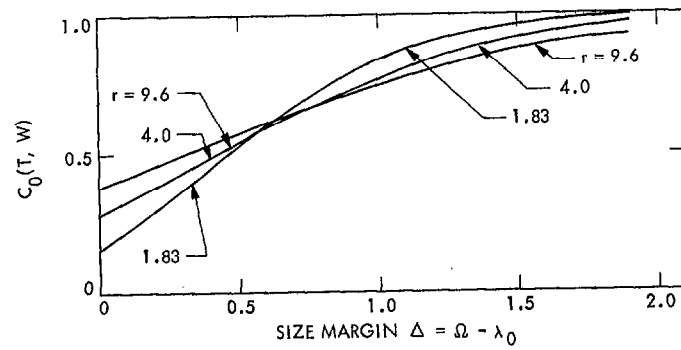
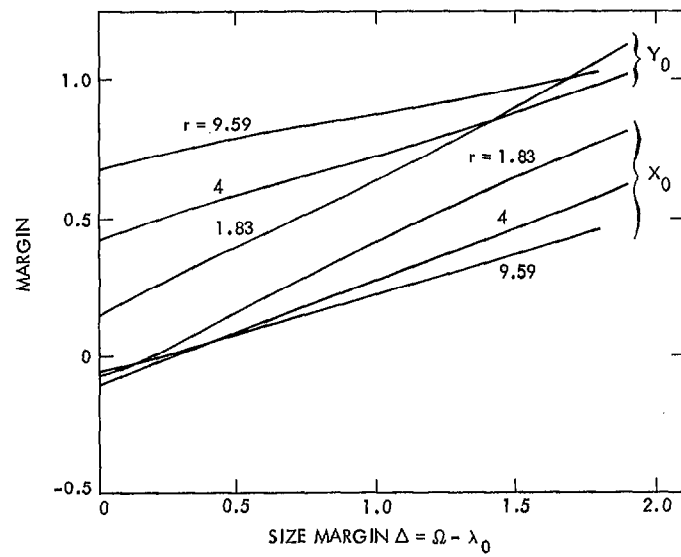


Fig. 6. Optimized duration and effort margins for given size margin  $\Delta$ , Walston-Felix data



**Fig. 7. Optimum project confidence factor for given size margin  $\Delta$ , Freburger-Basili data**



**Fig. 8. Optimized duration and effort margins for given size margin  $\Delta$ , Freburger-Basili data**